

Better bitmap performance with Roaring bitmaps

Samy Chambi^a, Daniel Lemire^{b,*}, Owen Kaser^c, Robert Godin^a

^aDépartement d'informatique, UQAM, 201, av. Président-Kennedy, Montreal, QC, H2X 3Y7 Canada

^bLICEF, Université du Québec, 5800 Saint-Denis, Montreal, QC, H2S 3L5 Canada

^cComputer Science and Applied Statistics, UNB Saint John, Saint John, NB, E2L 4L5 Canada

Abstract

Bitmap indexes are commonly used in databases and search engines. By exploiting bit-level parallelism, they can significantly accelerate queries. However, they can use much memory. Thus we might prefer compressed bitmap indexes. Following Oracle's lead, bitmaps are often compressed using run-length encoding (RLE). In this work, we introduce a new form of compressed bitmaps called *Roaring*, which uses packed arrays for compression instead of RLE. We compare it to two high-performance RLE-based bitmap encoding techniques: WAH (Word Aligned Hybrid compression scheme) and Concise (Compressed 'n' Composable Integer Set). On synthetic and real data, we find that Roaring bitmaps (1) often compress significantly better (e.g., $2\times$) and (2) are faster than the compressed alternatives (up to $900\times$ faster for intersections).

Keywords: bitmap index, compression, indexing

1. Introduction

A bitmap (or bitset) is a binary array that we can view as an efficient and compact representation of an integer set. Given a bitmap of n bits, the i^{th} bit is set to one if the i^{th} integer in the range $[0, n - 1]$ exists in the set. For example, the sets $\{3, 4, 7\}$ and $\{4, 5, 7\}$ might be stored in binary form as 10011000 and 10110000. We can compute the union or the intersection between two such corresponding lists using bitwise operations (OR, AND) on the bitmaps (e.g., 10111000 and 10010000 in our case).

*Corresponding author. Tel.: 00+1+ 514 843-2015 ext. 2835; fax: 00+1+514 843-2160.

Email addresses: `chambi.samy@gmail.com` (Samy Chambi), `lemire@gmail.com` (Daniel Lemire), `o.kaser@computer.org` (Owen Kaser), `godin.robert@uqam.ca` (Robert Godin)

Bitmaps are part of the Java platform (`java.util.BitSet`).

When the cardinality of the set $|S|$ is relatively large compared to the universe size, n (e.g., $|S| > n/64$ on 64-bit processors), bitmaps are often superior to other comparable data structures such as arrays, hash sets or trees. However, on moderately low density bitmaps ($n/10000 < |S| < n/64$), compressed bitmaps such as Concise can be preferable [1].

Most of the recently proposed compressed bitmap formats are derivative from Oracle’s BBC [2]: WAH [3], Concise [1], EWAH [4], COMPAX [5], VLC [6], VAL-WAH [7], etc. Wu et al.’s WAH is probably the best known. WAH divides a bitmap of n bits into $\lceil \frac{n}{w-1} \rceil$ words of $w-1$ bits, where w is a convenient word length (e.g., $w = 32$). WAH distinguishes between two types of words: words made of just $w-1$ ones ($11\cdots 1$) or just $w-1$ zeros ($00\cdots 0$), are *fill words*, whereas words containing a mix of zeros and words (e.g., $101110\cdots 1$) are *literal words*. Literal words are stored using w bits: the most significant bit is set to zero and the remaining bits store the heterogeneous $w-1$ bits. Sequences of homogeneous fill words (all ones or all zeros) are also stored using w bits: the most significant bit is set to 1, the second most significant bit indicates the bit value of the homogeneous block sequence, while the remaining $w-2$ bits store the run length of the homogeneous block sequence.

When compressing a sparse bitmap, e.g., corresponding to the set $\{0, 2(w-1), 4(w-1), \dots\}$, WAH can use $2w$ bits per set bit. Concise reduces this memory usage by half [1]. It uses a similar format except for coded fill words. Instead of storing the run length r using $w-1$ bits, Concise uses only $w-2-\lceil \log_2(w) \rceil$ bits, setting aside $\lceil \log_2(w) \rceil$ bits as *position* bits. These $\lceil \log_2(w) \rceil$ *position* bits encode a number $p \in [0, w)$. When $p = 0$, we decode $r+1$ fill words. When it is non-zero, we decode r fill words preceded by a word that has its $(p-1)^{\text{th}}$ bit flipped compared to the following fill words.

Consider the case where $w = 32$. Concise can code the set $\{0, 62, 124, \dots\}$ using only 32 bits/integer, in contrast to WAH which requires 64 bits/integer.

Though they reduce memory usage, these formats derived from BBC have slow random access compared to an uncompressed bitmap. That is, checking or changing the i^{th} bit value is an $O(n)$ -time operation. Thus, though they represent an integer set,

we cannot quickly check whether an integer is in the set. This makes them unsuitable for some applications [8]. Moreover, RLE formats have a limited ability to quickly skip data. For example, suppose that we are computing the bitwise AND between two compressed bitmaps. If one bitmap has long runs of zeros, we might wish to skip over the corresponding words in the other bitmap. Without an auxiliary index, this might simply be impossible with formats like WAH and Concise.

Instead of using RLE and sacrificing random access, we propose to partition the space $[0, n)$ into *chunks* and to store dense and sparse chunks differently [9]. On this basis, we introduce a new bitmap compression scheme called *Roaring*. Roaring bitmaps store 32-bit integers in a compact and efficient two-level indexing data structure. In our example $(\{0, 62, 124, \dots\})$, it would use only ≈ 16 bits/integer, half of Concise’s memory usage. Moreover, on the synthetic-data test proposed by Colantonio and Di Pietro [1], it is at least four times faster than WAH and Concise. In some instances, it can be hundreds of times faster.

2. Roaring bitmap

We partition the range of 32-bit indexes $([0, n))$ into chunks of 2^{16} integers sharing the same 16 most significant digits. We use specialized containers to store their 16 least significant bits.

When a chunk contains no more than 4096 integers, we use a sorted array of packed 16-bit integers. When there are more than 4096 integers, we use a 2^{16} -bit bitmap. Thus, we have two types of containers: an array container for *sparse* chunks and a bitmap container for *dense* chunks. The 4096 threshold insures that at the level of the containers, each integer uses no more than 16 bits: we either use 2^{16} bits for more than 4096 integers, using less than 16 bits/integer, or else we use exactly 16 bits/integer.

The containers are stored in a dynamic array with the shared 16 most-significant bits: this serves as a first-level index. The array keeps the containers sorted by the 16 most-significant bits. We expect this first-level index to be typically small: when $n = 1\,000\,000$, it contains at most 16 entries. Thus it should often remain in the CPU cache. The containers themselves should never use much more than 8 kB.

Each container keeps track of its cardinality (number of integers) using a counter. Thus computing the cardinality of a Roaring bitmap can be done quickly: it suffices to sum at most $\lceil n/2^{16} \rceil$ counters. It also makes it possible to support rank and select queries faster than with a typical bitmap: rank queries count the number of set bits in a range $[0, i]$ whereas select queries seek the location of the i^{th} set bit.

The overhead due to the containers and the dynamic array means that our memory usage can exceed 16 bits/integer. However, as long as the number of containers is small compared to the total number of integers, we should never use much more than 16 bits/integer. We assume that there are far fewer containers than integers.

3. Access operations

To check for the presence of a 32-bit integer x , we first seek the container corresponding to $x/2^{16}$, using binary search. If a bitmap container is found, we access the $(x \bmod 2^{16})^{\text{th}}$ bit. If an array container is found, we use a binary search again.

When removing an integer, a bitmap container might become an array container if its cardinality reaches 4096. When adding an integer, an array container might become a bitmap container when its cardinality exceeds 4096. When this happens, a new container is created with the updated data while the old container is discarded.

4. Logical operations

We implemented various operations on Roaring bitmaps, including union (bitwise OR) and intersection (bitwise AND). A bitwise operation between two Roaring bitmaps consists of iterating and comparing the 16 high-bits integers (keys) on the first-level indexes. For better performance, we maintain sorted first-level arrays. Two keys are compared at each iteration. On equality, a second-level logical operation between the corresponding containers is performed. This always generates a new container. If the container is not empty, it is added to the result along with the common key. Then iterators positioned over the first-level arrays are incremented by one. When two keys are not equal, the array containing the smallest one is incremented by one position, and if a union is performed, the lowest key and a copy of the corresponding container are

added to the answer. When computing unions, we repeat until the two first-level arrays are exhausted. And when computing intersections, we terminate as soon as one array is exhausted.

Sorted first-level arrays allows first-level comparisons in $O(n_1 + n_2)$ time, where n_1 and n_2 are the respective lengths of the two compared arrays. We also maintain the array containers sorted for the same advantages. As containers can be represented with two different data structures, bitmaps and arrays, a logical union or intersection between two containers involves one of the three following scenarios:

Bitmap vs Bitmap: We iterate over 1024 64-bit words. For unions, we simply perform 1024 bitwise ORs and write the result to a new bitmap container. The resulting cardinality is computed efficiently in Java using the `Long.bitCount` method. More care is needed for intersections. First, we compute the cardinality of the result, using 1024 bitwise AND instructions. If the cardinality is larger than 4096, then we proceed as with the union, writing the result of bitwise ANDs to a new bitmap container. Otherwise, we create a new array container. We extract the set bits from the bitwise ANDs on the fly, using a fast method such as Java's `Long.numberOfTrailingZeros`.

Bitmap vs Array: When one of the two containers is a bitmap and the other one is a sorted dynamic array, the intersection can be computed very quickly: we iterate over the sorted dynamic array, and verify the existence of each 16-bit integer in the bitmap container. The result is written out to an array container. Unions are also efficient: we create a copy of the bitmap and simply iterate over the array, setting the corresponding bits.

Array vs Array: For unions, if the sum of the cardinalities is less than 4096, we use a merge algorithm between the two arrays. Otherwise, we set the bits corresponding to both arrays in a bitmap container. If the cardinality is no more than 4096, we convert the bitmap container. For intersections, we use a simple merge (akin to what is done in merge sort) when the two arrays have cardinalities that differ by less than a factor of 64. Otherwise, we use *galloping* intersections (e.g., see [8]). The result is always written to a new array container.

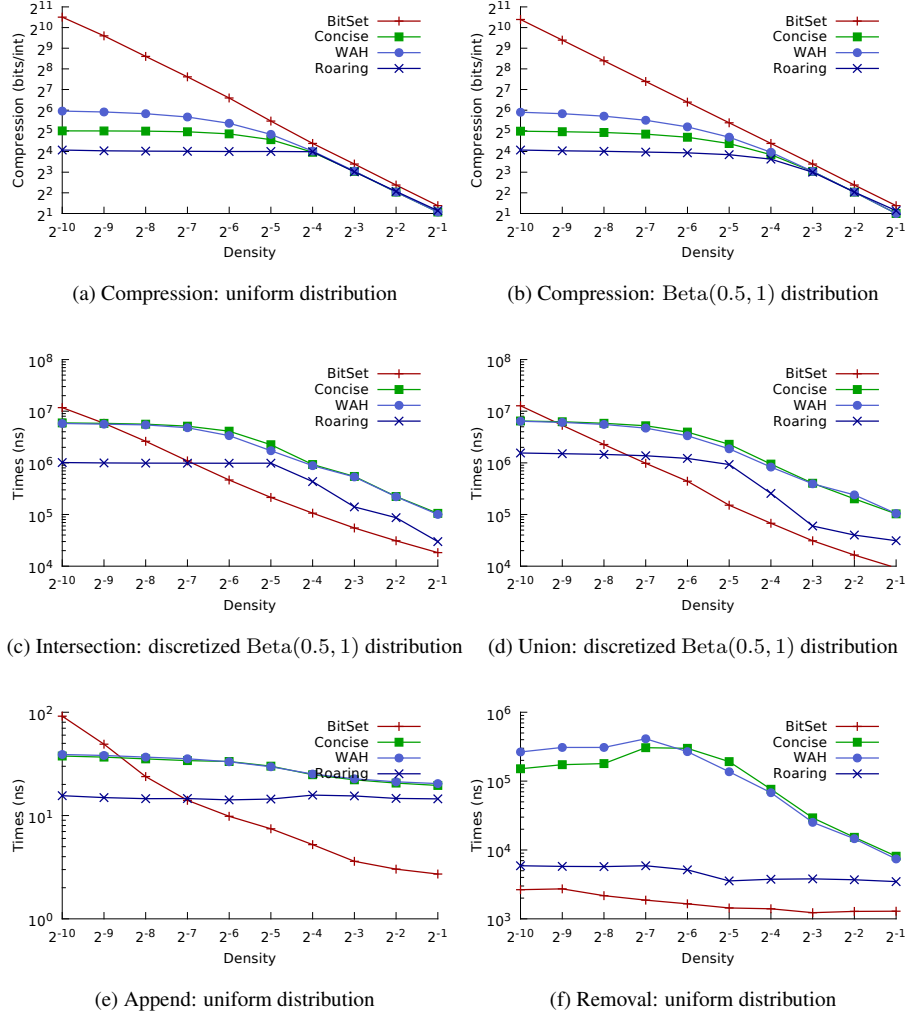


Figure 1: Times and compression measurements: average of 100 runs

We could also execute these operations *in place* when needed. That is, instead of generating a new bitmap container when computing the union between two bitmap containers, we could modify one of the two containers. It is comparatively more difficult to support in place operations with WAH and Concise.

5. Experiments

We performed a series of experiments to compare the time-space performance of Roaring bitmaps with the performance of other well-known bitmap indexing schemes: Java’s `BitSet`, WAH 32-bit and Concise 32-bit. We used the CONCISE Java library for WAH and Concise (version 2.2). Our Roaring-bitmap implementation code is freely available at <http://roaringbitmap.org/>.

Benchmarks were performed on an AMD FX™-8150 eight-core processor running at 3.60 GHz and having 16 GB of RAM. We used the Oracle 64-Bit Server JVM version 1.7 on Linux Ubuntu 12.04.1 LTS.

To account for the just-in-time compiler in Java, we first run tests without recording the timings. Then we repeat the tests several times and report an average.

5.1. Synthetic experiments

We began by reproducing Colantonio and Di Pietro’s synthetic experiments [1]. However, while they included alternative data structures such as Java’s `HashSet`, we focus solely on bitmap formats for simplicity. Our results are generally consistent with Colantonio and Di Pietro’s results given the fact that we have a better processor.

Data sets of 10^5 integers were generated according to two synthetic data distributions: uniform and discretized Beta(0.5, 1) distributions. (Colantonio and Di Pietro described the latter as a *Zipfian* distribution.) The four schemes were compared on several densities d varying from 2^{-10} to 0.5. To generate an integer, we first picked a floating-point number y pseudo-randomly in $[0, 1)$. When we desired a uniform distribution, we added $\lfloor y \times \text{max} \rfloor$ to the set. In the *beta*-distributed case, we added $\lfloor y^2 \times \text{max} \rfloor$. The value max represents the ratio between the total number of integers to generate and the desired density (d) of the set, i.e.: $\text{max} = 10^5/d$. Because results on uniform and Beta(0.5, 1) distributions are often similar, we do not systematically present both.

Figs. 1a and 1b show the average number of bits used by Java’s `BitSet` and the three bitmap compression techniques to store an integer in a set. On sparse bitmaps, Roaring bitmaps require 50 % of the space consumed by Concise and 25 % of WAH space.

We also report on intersection and union times. That is, we take two bitmaps and generate a new bitmap representing the intersection or union. For the `BitSet`, it means that we first need to create a copy (using the `clone` method) since bitwise operations are in-place. Figs. 1c and 1d present the average time in nanoseconds to perform intersections and unions between two sets of integers. Roaring bitmaps are $\times 4 - \times 5$ times faster than Concise and WAH for intersections on all tested densities. Results for unions are similar except that for moderate densities ($2^{-5} \leq d \leq 2^{-4}$), Roaring is only moderately (30 %) faster than Concise and WAH. `BitSet` outperforms the other schemes on dense data, but it is $> 10\times$ slower on sparse bitmaps.

We measured times required by each scheme to add a single element a to a sorted set S of integers, i.e.: $\forall i \in S : a > i$. Fig. 1e shows that Roaring requires lower times than WAH and Concise. Moreover, WAH and Concise do not support the efficient insertion of values in random order, unlike Roaring bitmaps. Finally, we measured the time needed to remove one randomly selected element from an integers set (Fig. 1f). We observe that Roaring bitmaps have much better results than the two other compressed formats.

5.2. Real-data experiments

Tables 1–2c presents results for the five real datasets used an earlier study of compressed bitmap indexes [10]. There are only two exceptions:

- We only use the September 1985 data for the WEATHER dataset (an approach others have used before [11]), which was otherwise too large for our test environment.
- We omitted the CENSUS2000 dataset because it contains only bitmaps having an average cardinality of 30 over a large universe ($n = 37\,019\,068$). It is an ill-suited scenario for bitmaps. Because of the structure overhead, Roaring bitmaps used $4\times$ as much memory as Concise bitmaps. Still, Roaring bitmaps were about $4\times$ faster when computing intersections.

The dataset are taken as-is: we do not sort them prior to indexing.

Table 1: Sampled bitmap characteristics and Roaring size.

	CENSUS1881	CENSUSINCOME	WIKILEAKS	WEATHER
Rows	4 277 807	199 523	1 178 559	1 015 367
Density	1.2×10^{-3}	1.7×10^{-1}	1.3×10^{-3}	6.4×10^{-2}
Bits/Item	18.7	2.92	22.3	5.83

For each dataset, a bitmap index was built. Then we chose 200 bitmaps from the index, using an approach similar to stratified sampling to control for the large range of attribute cardinalities. We first sampled 200 attributes, with replacement. For each sampled attribute, we selected one of its bitmaps uniformly at random. The 200 bitmaps were used as 100 pairs of inputs for 100 pairwise ANDs and ORs; Tables 2b–2c show the time factor increase if Roaring is replaced by `BitSet`, WAH or Concise. (Values below 1.0 indicate cases where Roaring was slower.) Table 2a shows the storage factor increase when Roaring is replaced by one of the other approaches.

Roaring bitmaps are always faster, on average, than WAH and Concise. On two datasets (CENSUS1881 and WIKILEAKS), Roaring bitmaps are faster than `BitSet` while using much less memory (40× less). On the two other datasets, `BitSet` is more than twice as fast as Roaring, but it also uses three times as much memory. When comparing the speed of `BitSet` and Roaring, consider that Roaring pre-computes the cardinality at a chunk level. Thus if we need the cardinality of the aggregated bitmaps, Roaring has the advantage. On the WIKILEAKS dataset, Concise and WAH offer better compression than Roaring (by about 30 %). This is due to the presence of long runs of ones ($11 \dots 1$ fill words) which Roaring does not compress.

Results on the CENSUS1881 dataset are striking: Roaring is up to 900× faster than the alternatives. This is due to the large differences in the cardinalities of the bitmaps. When intersecting a sparse bitmap with a dense one, Roaring is particularly efficient.

Table 2: Results on real data

(a) Size expansion if Roaring is replaced with other schemes.

	CENSUS1881	CENSUSINCOME	WIKILEAKS	WEATHER
Concise	2.2	1.4	0.79	1.4
WAH	2.4	1.6	0.79	1.5
BitSet	42	2.9	55	3.5

(b) Time increase, for AND, if Roaring is replaced with other schemes.

	CENSUS1881	CENSUSINCOME	WIKILEAKS	WEATHER
Concise	920	6.6	8.3	6.3
WAH	840	5.9	8.2	5.4
BitSet	730	0.42	28	0.64

(c) Time increases, for OR, if Roaring is replaced with other schemes.

	CENSUS1881	CENSUSINCOME	WIKILEAKS	WEATHER
Concise	34	5.4	2.1	3.9
WAH	31	4.9	2.1	3.4
BitSet	29	0.43	6.7	0.48

6. Conclusion

In this paper, we introduced a new bitmap compression scheme called Roaring. It stores bitmap set entries as 32-bit integers in a concise two-level indexes. In comparison with two competitive bitmap compression schemes, WAH and Concise, Roaring uses less memory and is faster.

In future work, we will consider improving the performance further. We might add new types of containers. In particular, we might make use of fast packing techniques to optimize the storage use of the array containers [12]. We could also extend Roaring bitmaps to 64-bit integers.

References

- [1] A. Colantonio, R. Di Pietro, Concise: Compressed 'n' composible integer set, *Inform. Process. Lett.* 110 (2010) 644–650.
- [2] G. Antoshenkov, Byte-aligned bitmap compression, in: *DCC'95*, IEEE Computer Society, Washington, DC, USA, 1995, p. 476.
- [3] K. Wu, K. Stockinger, A. Shoshani, Breaking the curse of cardinality on bitmap indexes, in: *SSDBM'08*, Springer, 2008, pp. 348–365.
- [4] D. Lemire, O. Kaser, K. Aouiche, Sorting improves word-aligned bitmap indexes, *Data Knowl. Eng.* 69 (2010) 3–28.
- [5] F. Fusco, M. P. Stoecklin, M. Vlachos, NET-FLi: On-the-fly compression, archiving and indexing of streaming network traffic, *Proc. VLDB Endow.* 3 (2010) 1382–1393.
- [6] F. Corrales, D. Chiu, J. Sawin, Variable length compression for bitmap indices, in: *DEXA'11*, Springer-Verlag, Berlin, Heidelberg, 2011, pp. 381–395.
- [7] G. Guzun, G. Canahuate, D. Chiu, J. Sawin, A tunable compression framework for bitmap indices, in: *ICDE'14*, IEEE, 2014.
- [8] J. S. Culpepper, A. Moffat, Efficient set intersection for inverted indexing, *ACM T. Inform. Syst.* 29 (2010) 1:1–1:25.
- [9] O. Kaser, D. Lemire, Attribute value reordering for efficient hybrid OLAP, *Inform. Sciences* 176 (2006) 2304–2336.
- [10] D. Lemire, O. Kaser, E. Gutarra, Reordering rows for better compression: Beyond the lexicographical order, *ACM T. Database Syst.* 37 (2012). Article 20.
- [11] K. Beyer, R. Ramakrishnan, Bottom-up computation of sparse and iceberg CUBEs, *SIGMOD Rec.* 28 (1999) 359–370.
- [12] D. Lemire, L. Boytsov, Decoding billions of integers per second through vectorization, *Softw. Pract. Exp.* (2013). In Press.